Possible architectures for an hybrid correlator

G. Comoretto[1]

[1]Osservatorio Astrofisico di Arcetri

Arcetri Technical Report N° 8/2000
Firenze, December 2000

Abstract

For the ALMA advanced correlator, an hybrid architecture has been proposed. By dividing the digitized band into several smaller sub-bands, using digital filtering, downconversion and Nyquist subsampling (decimation), one can achieve a significant reduction in the size of the correlator. Various solutions have been proposed to implement this concept. In this repost these solutons are compared, in term of performances and of required hardware, both for the signal conditioning part (antenna based) and for the correlator.

1 Introduction

In the meeting held at Bordeaux Observatory on 18-19 April 2000, a design for an *advanced correlator* for ALMA has been proposed [4]. This design is based on the concept of a hybrid XF correlator. The digitized signal is first separated into several subbands, using digital filters and either digital downconverters or Nyquist subsampling. Correlation of subbands then requires much less correlator computing power than for a traditional time-multiplexed correlator, at the expense of a far more complex signal conditioning electronics after the sampler. Since the latter is proportional to the number of antennas N, while correlator size is proportional to N^2, an hybrid correlator becomes advantageous for large N interferometers.

Several problems have been raised for an hybrid correlator. Multiple quantization may degrade SNR. Subbands have different fringe frequency. Subbands alignment (platforming) may be problematic, due to differences in quantization effects among subbands. Limited number of channels in each subband increases edge effects (e.g. problems of determining phase for channel 0).

To be competitive with respect to the currently proposed baseline correlator, the hybrid design should not require an increased number of interconnection between antenna units and correlator units. It should offer comparable or better flexibility in trading resolution for bandwidth and/or number of polarizations, and either a sensible decrease in overall cost, increased performance, or both.

In this report, I will compare several possible architectures for an hybrid correlator, in term of performances and complexity, and examine the impact of these architectures on the problems raised.

This report is a contribution to the European ALMA Back-End Electronics Team. Research for ALMA in Italy is supported by a grant of the Italian CNAA.

2 Problems and advantages of a hybrid correlator

A hybrid correlator has intrinsically some problems with respect to an analog time multiplexed counterpart. These problems are in part independent from the particular architecture, and are therefore considered separately.

2.1 Dynamic range

In a digital system quantization introduces possibilities for sever nonlinearities. Digital LO is simulated with a stepwise sinusoid, introducing high order harmonics that may fall back in the observed band. Two and three bit quantization requires a linearization that cannot be applied if the signal is applied to a digital filter, Requantization after filtering introduces further nonlinearities. Digital filtering must reject with high accuracy the portion of the input spectrum not specifically in the passband, to prevent aliasing of strong out-of-band signals.

The final effect of all these factors are a reduction of signal to noise ratio, that can in principle be tested by simulation, and a reduction in dynamic range. When a strong spectral feature is present, intermodulation products and limited rejection of undesired frequencies produce ghosts at different frequencies, over the entire IF band.

Some of these effects can be quantified analytically. Filter rejection depends mainly on the number of bits used for the coefficients. With 8 bits, a rejection of 40 dB can be obtained, resulting in a dynamic range of 1:10,000. To obtain a dynamic range of 38 dB, at least 4 bits must be used in the representation of the LO sinusoid (see [3]). It is important to note that, in the limit of very low frequency LO's, as used for example in fringe rotation, LO harmonics occur at very low frequencies, producing a (negligible) broadening of any spectral feature but not ghost images. The presence of ghost images can be a serious issue for architectures with independently tunable subbands, if more than 38 dB of ghost rejection is required.

2.2 Differential fringe stopping

In ALMA, fringe stopping is done in the LO chain. Residual fringe phase across the IF bandwidth may require additional fringe rotation in each subband, depending on the accuracy of the delay tracking. In the baseline correlator, delay tracking is done with an accuracy of $\pm 1/2$ sampler clock. If in the hybrid correlator delay tracking is done only after the final resampling, with an accuracy of one subband clock sample (4 ns for 250 MHz) an independent fringe rotator is required for each subband.

However, in an hybrid concept, it is very easy to obtain fractional bit delay. First, a barrel shifter after the sampler can provide fine delay with accuracy of one sampler clock (250 ps). If fringe frequency is chosen for midband, a maximum phase error of $\pm\pi/4$ occurs at the two extreme subbands (phase is identically zero at midband and oscillates between $\pm\pi/4$ at frequency 0 and $f_s/2$). This causes a signal loss of about 19% for these subbands.

Using a variable phase sampling clock it is possible to achieve fractional delay with unlimited accuracy. Only one sampling clock is required for each antenna, since delay depends on the geometric delay and not on sky frequency. Phase adjustment can be done either continuously (e.g. using a DDS for sampling clock generation) or discretely.

Alternatively, one can implement fine delay in the digital filters, in principle with infinite resolution. Marc Torres extimates that to achieve 1 degree phase errors, filter coefficients need to be updated every 600μs. This holds also for fractional delay implemented in the sampling clock.

Using separate (digital) local oscillators for each subband, it is possible to adjust the fringe rate and stop them differentially in each subband. If integral delay compensation is done at the sampler, phase slope is limited to $\pi/2N$ in each of the N subbands, plus a residual phase drift that can be compensated by the digital LO. Digital local oscillators have been proposed in several instantations of a hybrid correlator, but there may be limitations in the clock frequency of different subbands, resulting in only limited phase tracking.

3 Comparison of different correlator schemes

Several possibilities have been proposed, either as concepts and as working instruments, for a hybrid correlator. I will try to analyze all of them, comparing their performances in terms of resources used for a given required configuration. I will assume as a reference configuration a single IF, with 2 GHz instantaneous bandwidth, correlation done at 250 MS/s, no polarizations, N antennas, 256 frequency points (i.e. resolution of 7.8 MHz with rectangular windowing).

In specifying lags for the correlators, I will count only positive lags. A total of N^2 correlators are required, to obtain N autocorrelations and $N(N-1)/2$ crosscorrelations, with negative lags for baseline (i, j) calculated in correlator (j, i). In some correlator architectures (Bos counterflow correlator), the crosscorrelation for positive and negative lags are performed in the same place. This does not affect the total number of channels, however, and is therefore completely outside the scope of this note.

In the comparison, I will assume fixed performance and compare the amount of necessary hardware. When the same performances can be achieved with reduced hardware, one can trade increased performances for smaller savings in hardware. However, since the scope of this report is to compare different architectures, no effort is made in determining the best cost/performance compromise.

3.1 Common part

Sampler, fiber link and associated control hardware are common to all architectures. This hardware is shown in fig.1.

The IF output signals are sampled at the antenna, with an optional fractional-bit delay. Fringe rate is corrected at 2nd LO. The signal is transmitted to the correlator using a fiber link.

At the moment it is not specified whether time multiplexing is done before or after transmission in the fiber link. Given the availability of fast data fiber links, it seems advisable to use commercial framing and multiplexing circuitry, and considering the link as a virtual connection capable of delivering the sampler outputs (4IF $\times$ 3bits $\times$ 4MS/s) to the correlator building, together with framing information.

Before processing, signal is parallelized to a convenient manageable clock rate. In the baseline correlator this is 125 MHz ($\times 32$ time multiplexing). It appears reasonable to assume that a 250 MHz clock rate can be managed, both for the PCB and for the CMOS chips, and therefore it will be assumed for this report. All architectures considered here assume an input composed of 16×3-bit samples, at a clock rate

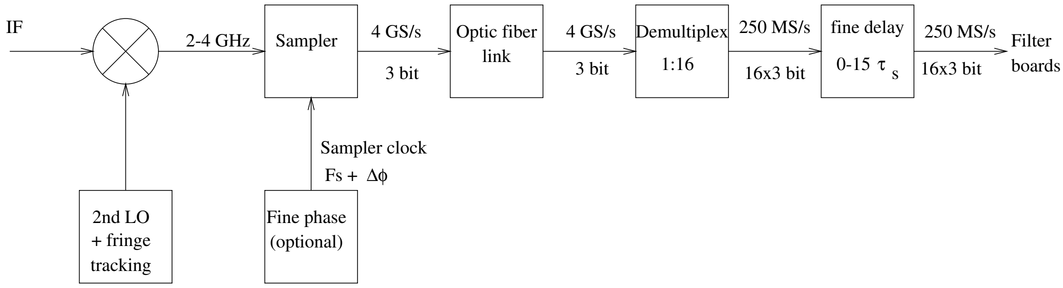

Figure 1: IF conversion, digitizing, and fine delay correction. These elements are common to all the examined architectures. Some elements are optional, i.e. their presence is not still decided. These include fine delay correction in the sampler (common to all samplers of a given antenna) and a barrel shifter for fine delay tracking.

of 250 MS/s, representing 16 consecutive samples at the correlator. Most considerations do not depend on the exact time multiplexing factor, however.

After parallelizing, it is straightforward to implement a fine delay with a range of 16 sampler clocks, and an accuracy of one clock sample. This reduces the phase error to $\pm\pi/4$ at the edges of the IF band, if the LO is programmed to compensate for fringe frequency at midband.

3.2 Baseline correlator

First, we consider an enhanced ALMA correlator, with the same architecture of the interim correlator, with faster and denser chips. This will serve as a base for comparison.

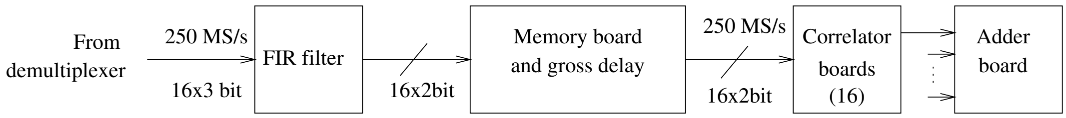

Figure 2: Baseline correlator architecture. Signal is filtered, for band selection, and time multiplexed in a memory buffer board, to be processed in parallel in N correlator boards

The correlator architecture is shown in fig. 2. A FIR filter is used for band selection only. The signal is then buffered in a memory bank, that performs also gross delay compensation. At the output of the memory board, the signal si available as 16 independent "time slots", that are separately correlated, and then summed together.

Each correlator must provide 256 lags, for a total of 256×16 lags per baseline, and 256 total frequency channels.

The memory board represents a significant amount of hardware. In the proposed interim correlator, it must provide storage for two banks of 1 ms of data, i.e. 8 Msamples.

The FIR filter must perform lowpass or bandpass filtering. Bandpass filtering is required for the smaller bandwidths, to avoid using the less accurate IF band edges. The total number of used filter taps is constant, since at lower bandwidths less parallel channels are needed. For a 90% usable bandwidth, the total number of taps (at 250 MHz filter clock) is 2048, independently from the final bandwidth (for bandwidths equal or greater to 125 MHz)

3.3 WIDAR design

This concept is described in [2], and is illustrated in fig. 3. The signal to be analyzed is separated into 16 contiguous subbands by FIR bandpass filters. Each subband is then resampled at 1/16 of the original frequency, ad thus aliased to baseband. The filter response is nominally flat in the whole subband, with an undesired spillover in the two adjacent ones. To remove the spillover, an offset (different for each antenna) is added to the second LO, and then removed after filtering and resampling. While the passband is stopped, the spillover is modulated at two times the offset, and rejected if the offset is chosen exactly sa a multiple of the elementary integration period. The unwanted spillover contributes however to the noise, that is increased by at most $\sqrt{2}$ in a small area at beginning and end of each subband.

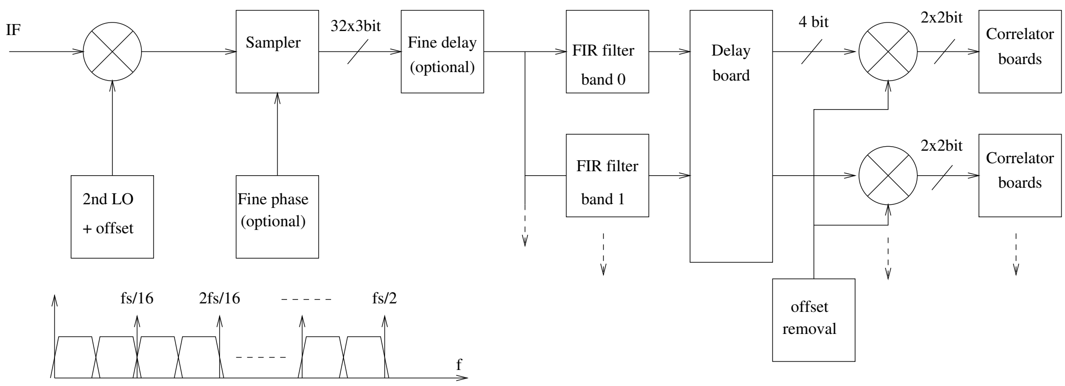

Figure 3: WIDAR correlator architecture. Signal is decomposed into N disjoint bands. Sideband rejection is done by adding a small offset frequency in the 2^{nd} LO, and removing it in a digital 3^{rd} LO, before complex correlation.

This concept is at the base of the existing correlator for the expanded VLA [references]

A delay board must still be used. However, the maximum delay is about 60 μs, or 15 KSample at the reduced sampling rate. Commercial FIFO chips exist for such memory depth.

Each correlator board processes only the corresponding subband. Therefore, a total of 256 lags are required for each baseline (16 subbands $\times$ 16 lags). This is the main advantage of an hybrid design.

The filters must have a much higher number of taps, in practise 16 independent filters with full length are required. The total number of taps is therefore 16 times those required for the baseline correlator.

If the filters can be reprogrammed on-the-fly, e.g. using dual bank coefficient memories, fine delay compensation can be implemented by shifting the (continuous) FIR function by the appropriate amount, in principle with infinite resolution. However, since separate LO's are available for each subband, phase tracking can be more easily implemented in these LO's, and the residual phase error due to delay mismatch is reduced by a factor of 16 with respect to the full-band correlator.

The third LO's must be complex, and complex correlation is required. This doubles the number of (real) correlation channels, and quadruples the number of multipliers with respect to a real correlator. The correlator produces a spectrum in which negative frequencies contain only noise, and are discarded. If the LO's are implemented in the antenna boards, instead than in the correlator chips, also the number of interconnects from the antenna units to the correlator units is doubled. A complex correlation offers the advantage of measuring the phase also for channel zero, a net advantage for a correlator divided into many small subbands.

The image suppression method described above does not work for autocorrelations. Autocorrelation spectra are unusable in the small regions affected by the aliased spillover, at the beginning and end of each subband.

3.4 Digital SSB design (both sidebands used)

This design has been proposed by B. Anderson, and is described in fig. 4.

The concept is similar to the previous case, but in this case the subband position is chosen by tuning a digital quadrature LO. The bandpass filter performs also a quadrature phase shift of the I and Q signals, implementing a digital equivalent of a SSB mixer. Each LO produces two subbands, that can independently positioned in the IF, changing the LO frequency and the filter shape.

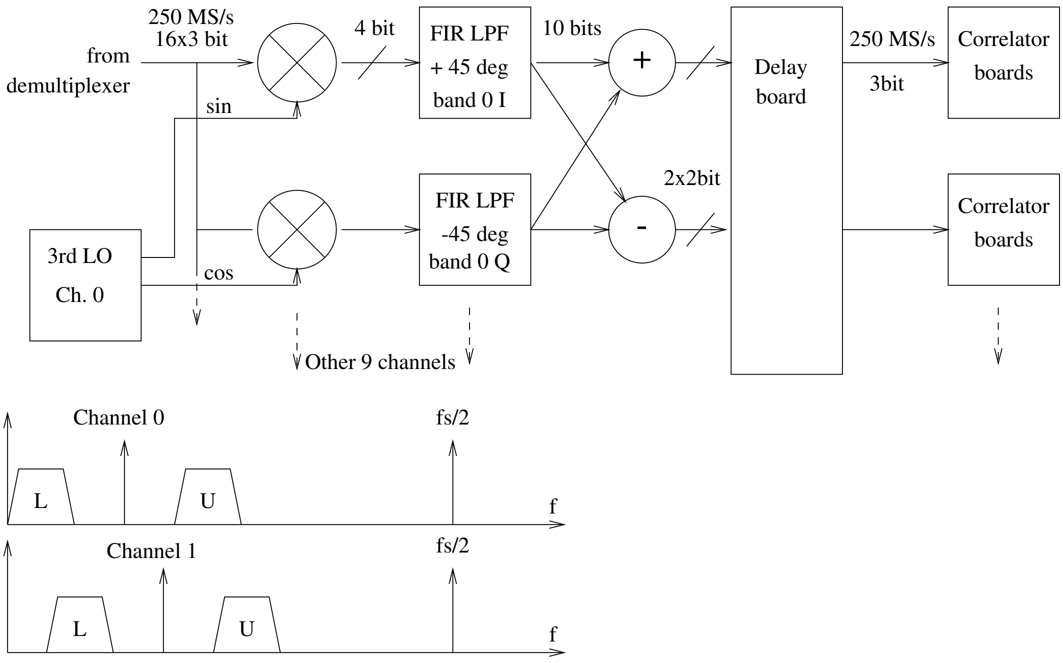

Figure 4: SSB correlator architecture. Signal is processed by $\approx N/2$ digital image-rejection SSB mixers. Upper and lower SSB are processed separately by conventional real correlators. A small overlap between adjacent subbands allow finite transition regions for the FIR filters. A total of at least $N/2+2$ converters ($N+4$ subbands) are used to cover the whole band.

The differential phase shift of $\pi/2$ between the two filters can be easily obtained imposing a $\pi/4$ phase shift to the response of one of the two filters, and adopting the time-reversed FIR response for the other. The FIR coefficients can be computed starting with a zero phase filter with the appropriate shape, applying the desired phase response in the frequency domain, backtransforming the resulting response, and truncating it to the initial FIR length. This rough procedure already gives a sideband rejection of more than 50 dB, and little stopband distortion. More accurate response can be obtained at this point by numerical optimization of the FIR coefficients. Bandpass filering is required to allow for a finite transition region in the phase response. If the sidebands are chosen adjacent, e.g. using a filter with bandpass starting near zero frequency, sideband rejection is poor (about 20 dB) in the lower portion of the band.

The filter passband width is chosen to be smaller than the final resampled bandwidth, and more subbands are required with respect to the previous case, to cover the whole IF band. The small overlap between adjacent subbands can be useful in alleviating the problem of subband stitching and calibration. A problem arises, however, from the fact that the upper and lower sidebands cannot be tuned independently, but must be separated by a multiple of the subsampling frequency, plus the filter guard band. Since the effective bandwidth and the correlator bandwidth of each channel are quite close, and therefore do not have a small common multiple, it is impossible to divide the total bandwidth into slightly overlapping subbands, *and* assign them to upper and lower channels of $(N/2+1)$ SSB's. At least $(N/2+2)$ SSB converters must be used, for example with the frequency coverage scheme shown in fig. 5. To minimize the problems of stitching together the subbands, it is advisable to tune the digital LO's to a multiple of the final channel separation.

A small offset can be added to each LO to partially compensate for differential fringe rate. Since sideband separation is about half the total bandwidth, phase error can be reduced by a factor of 2. This gives a maximum phase error of $\pm\pi/8$ at band edge, reducing the loss in SNR to 2.6%. This applies to almost all subbands, since the two sidebands are always separate by a significant fraction of the IF bandwidth. Phase tracking (fractional bit delay) can be implemented in the FIR filters, as described in chapter 2.2, basically shifting in time the continuous representation of the FIR function by a fraction of a tap, in both filters.

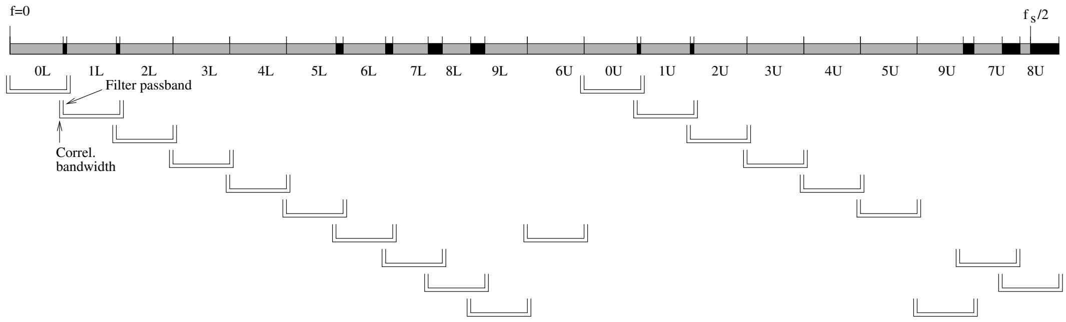

Figure 5: Frequency repartition for a system with 10 digital SSB converters. The overlapping regions are shown in black. The effective passband of each filter is 8/9 of the correlator bandwidth, i.e. 111 MHz.

Filter requirements are less stringent than for the baseline or the WIDAR correlator. For example, with the frequency coverage of this example, a passband of 8/9 of the correlator bandwidth can be used, leaving 1/9 of the correlator bandwidth as guard region (half of the guard region is folded back by aliasing, but does not fall in the useful part of the band). A reasonable FIR filter can be designed to meet these specifications with 512 taps. An example of a DSB passband using 512 taps, 1/16 IF bandwidth, 8 bit FIR coefficients (7 bit magnitude plus sign), 0.15 dB ripple, and an usable bandwidth of 8/9 of the correlator bandwidth is shown in fig. 6. Rejection of the unwanted sideband is limited only by truncation in the FIR coefficients.

Increasing the number of SSB converters relaxes the constrains on the filter, allowing it to be designed with fewer total taps. In fig. 7 a possible coverage using 11 converters, and filters with a passband of 100 MHz (256 taps each) is shown. This has the further advantage of increasing the overlap of adjacent subbands, but increases also the number of correlator units required.

The total number of correlation lags is slightly higher than for a WIDAR correlator, but they are all real, no complex correlation is correlator. This translates in a much simpler correlator, with a reduction of almost a factor of 4 in the total number of gates, despites the increased number of correlation planes required.

The first and last spectral channels in each subband fall in the transition band of the filter, and have undetermined phase. Therefore they must be discarded, and this constrains the *minimum* number of lags in each correlator plane. For example, using 9 digital SSB's (18 correlators for the full band), at least 16 lags must be provided. Discarding 2 spectral points form each subspectrum, one obtains $14 \times 18 = 252$ points for the full bandwidth. In practise, one will have much more than 16 lags per subband, and the problem does not arise.

This configuration adds a lot of flexibility for the observations. For example, one can choose the subbands in order to observe specific transitions with higher resolution.

3.5 Digital SSB design (one sideband used)

A more complex, but more flexible configuration can be obtained by using only one sideband in each SSB converter. In this case, one increases by a factor of 2 the antenna hardware, but the subbands can be placed without constrains in the IF band.

The use of fully independently tunable subbands give some advantage with respect to the previous case. Subbands can be evenly spaced in the full band, requiring less SSB converters, and, more importantly, less correlator units. Flexibility is increased with respect to the previous case, although this is not so useful, since for more than $N/2$ subbands (that are always available) the full bandwidth will be used in any case. Increased flexibility is useful if one has several very narrow windows. In this case, the correlator planes are clocked at reduced speed, to increase resolution, and it might be useful, in some applications, to independently tune N subbands.

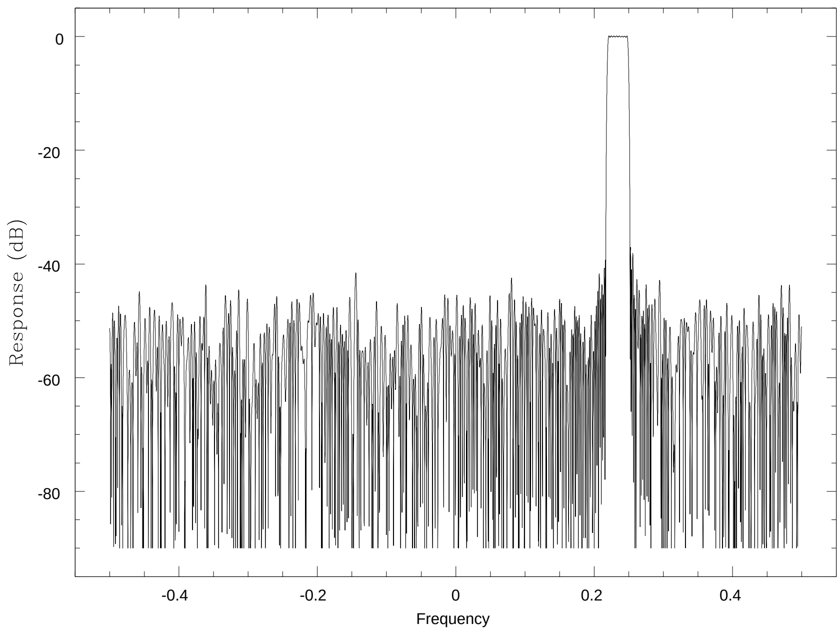

Figure 6: Frequency response for a single channel of a SSB converter. Only the filter and quadrature phase network are considered, using 512 taps, 8 bit FIR filters for the two I and Q channels. Bandwidth is expressed in term of the sampler frequency.

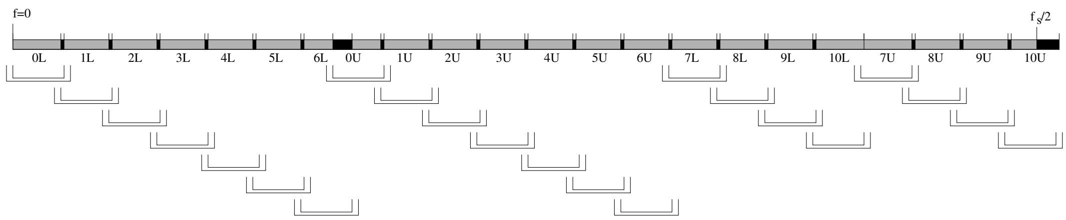

Figure 7: Frequency repartition for a system with 11 digital SSB converters. The overlapping regions are shown in black. The effective passband of each filter is 4/5 of the correlator bandwidth, i.e. 100 MHz. Subbands are much more evenly spaced than in the previous case.

As for the WIDAR correlator, partial (but almost complete) compensation for differential fringe rate can be achieved using the digital LO's. The residual phase error is $\pm\pi/4N$, with a negligible sensitivity loss.

3.6 SSB separation by complex correlation

A slightly more complex design can be obtained using a set of complex LO's/correlators (fig. 8).

Using a digital quadrature LO, ad a set of low pass filters with bandwidth equal to half the subband width, one obtain a complex signal with half sample rate with respect to a real signal. The number of interconnects to the correlator is thus the same than in the previous case. For example, 125 MS/s I and Q samples can be multiplexed on the same connection, at 250 MHz sample rate, for a total (real) bandwidth of 125 MHz (-62.5 to 6.25 MHz).

A complex correlator then separates the two sidebands, giving a seamless spectrum spanning the full subband. If the digital LO has a frequency greater than half the width of the subband, positive and negative frequency channels are independent, and a complex correlator with N lags is fully equivalent to a digital SSB followed by a real correlator with 2N lags. In particular, the correlator provides correct

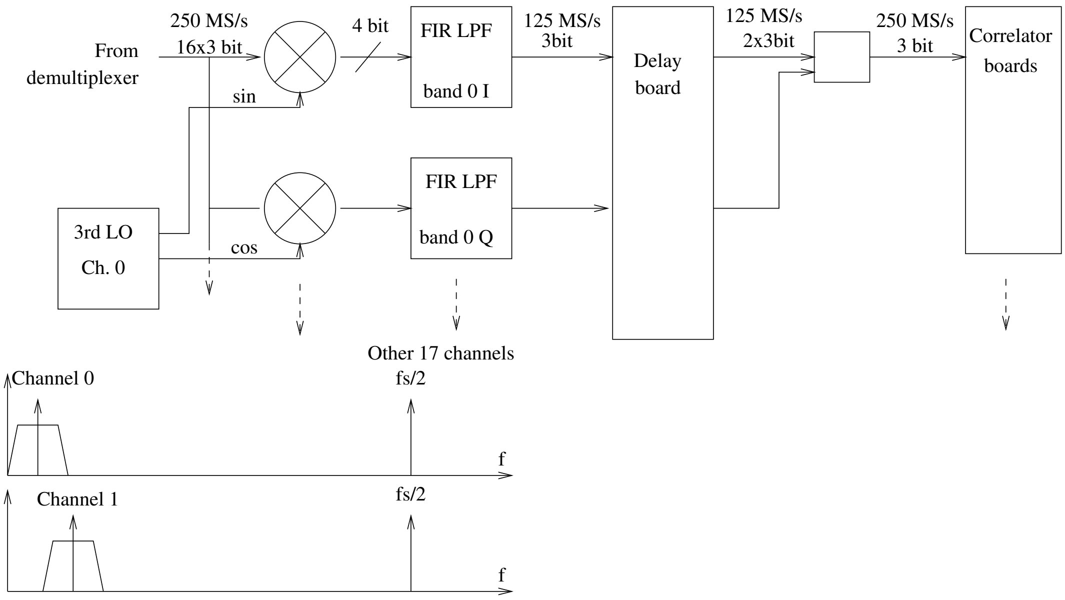

Figure 8: Complex SSB correlator architecture. Signal is split into $N+2$ complex subbands, processed by complex correlators. Number of LO's and filters is doubled with respect to a SSB correlator. Each subband of width B is represented by a complex stream with a data rate of B samples/s.

phase information for spectral channel at zero frequency.

The number of filters used is the same than in the previous case. Filter length is also the same, because low-pass instead of band-pass filtering is required, but bandwidth is halved.

The complex correlators operate now at half speed, and their complexity is comparable to a real correlator with nominal speed. Total number of real lags is the same than in the previous case (e.g. 8 complex lags, equivalent to 16 real lags, giving 16 spectral points). As in the previous case, extreme frequency points must be discarded.

Considering that the complexity of the filter board is exactly the same than for a digital SSB in which only one of the two sidebands is used, the use of a complex correlator does not appear justified. Moreover, a complex correlator operating at reduced precision (2 or 3 bit) may have nonlinearity problems, resulting in a nonideal sideband separation.

3.7 Summary

In table 3.7 various resources needed to implement a correlator with the different architectures are listed. The correlator bandwidth is set to 2 GHz, and the resolution to 256 spectral channels. Resources shown are those required for a single antenna and baseline (positive delays only, N^2 baselines required for full correlator).

WIDAR correlator cell is implemented as 2×2 multipliers, with parallel real and imaginary lags. Therefore it requires four multipliers and two shiftregisters per frequency point.

The complex BBC operates at half frequency, and two multipliers operating at full frequency can compute the complex product at the required speed. Real and imaginary samples are multiplexed in the same shiftregister. Since positive and negative frequencies represent independent frequency points, it requires one multiplier and one shiftregister cell per frequency point.

The three SSB configurations are very similar, differing only in the number of parallel filter units used. The complex correlator requires some extra circuitry, and therefore the last option requires in fact somewhat more complex circuitry, without any increase in performance with respect to the 18 SSB option. It might prove advantageous if complex correlation is advisable for other reasons.

A digital SSB, with different tradeoffs between correlator and filter complexity, appears a more con-

Architecture	FIR taps	Delay memory	Interconn. bandwidth	Shiftreg. cells	Multipl. (real)	$\Delta\phi$ track in LO
Scaling	N	N	N	N^2	N^2	
Baseline	1024	8 Ms	4 Gs/s	4096	4096	No
WIDAR	16384	256 Ks	4 Gs/s	512	1024	Yes
SSB (10 converters)	10240	320 Ks	5.0 Gs/s	320	320	Little
SSB (11 converters)	5632	352 Ks	5.5 Gs/s	352	352	Little
SSB (18 single output)	18432	288 Ks	4.5 Gs/s	288	288	Yes
Complex BBC	18432	288 Ks	4.5 Gs/s	288	288	Yes

Table 1: Comparison of resources used in different architectures to obtain 2 GHz BW, 256 spectral channels

venient architecture with respect to a WIDAR type design. Using only one sideband of the SSB may be advantageous in term of increased flexibility, ease of use, and differential fringe tracking, at the expense of a more complex antenna hardware. The exact number of subbands depends somewhat on the cost function of both the antenna and correlator units. More subbands require less critical filters in the antenna units, but more correlator planes.

Saving in correlator lags for an advanced correlator can be exploited either in complexity saving, increased performance (i.e. more spectral channels with the same number of lags), or both. For example, one may have 64 channels per correlator plane, for a total of 1152 lags per IF (roughly 1/4 with respect to the baseline correlator), still gaining a factor of 4 in spectral resolution.

References

[1] Anderson, B.: "FIR Filter Designs for ALMA" web page `http://www.jb.man.ac.uk/b̃a/firf.html` (1999)

[2] Carlson B. R., Dewdney P. E.: "Efficient wideband digital correlation", *Electronic Letters*, (2000)

[3] Comoretto G.: "A digital BBC for the ALMA interferometer", *ALMA memo series* n. 305 (2000)

[4] A. Baudry, M. de Vos eds. "Proceedings of the ALMA Backend & Correlator Subsystems meeting on Future Correlator Concept and Digital Filtering", Bordeaux, 18-19 April 2000.

Contents

1 Introduction **1**

2 Problems and advantages of a hybrid correlator **1**
2.1 Dynamic range 1
2.2 Differential fringe stopping 1

3 Comparison of different correlator schemes **2**
3.1 Common part 2
3.2 Baseline correlator 3
3.3 WIDAR design 3
3.4 Digital SSB design (both sidebands used) 4
3.5 Digital SSB design (one sideband used) 6
3.6 SSB separation by complex correlation 7
3.7 Summary 8